STIC-Biotech/ChemLib

145042

From: Chan, Christina
Sent: Monday, February 14, 2005 9:33 AM
To: Sullivan, Daniel; STIC-Biotech/ChemLib
Subject: RE: Rush sequence search 08/876132

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----
From: Sullivan, Daniel
Sent: Sunday, February 13, 2005 9:56 AM
To: Chan, Christina
Subject: Rush sequence search 08/876132

Hi Chris, Please approve this search as a rush for an after final amended case. Thanks.

Please search for the following in the pending and issued patent databases:

A nucleic acid comprising SEQ ID NO: 1;
A nucleic acid comprising SEQ ID NO: 2;
A nucleic acid encoding SEQ ID NO: 3.

Thank you.

Daniel M. Sullivan
Examiner AU 1636
Remsen Bldg.
Room 2A74

Tel: (571) 272-0779

Mailbox: 2C70

****
STAFF USE ONLY

Searcher:_____*Noble*_____
Searcher Phone: 2-
Date Searcher Picked up:___2/16/05
Date Completed:_____2/16/05
Searcher Prep/Rev. Time:___10
Online Time:_____ 10

****
Type of Search 3
NA Sequence: #_____3___
AA Sequence :#_____
Structure: #_____
Bibliographic:_____
Litigation:_____
Patent Family:_____
Other:_____

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM: *Compu(er*
WWW/Internet:_____
Other(Specify):_____

OM protein - nucleic search, using frame_plus_p2n model

Copyright (c) 1993 - 2005 Compugen Ltd.
GenCore version 5.1.6

Run on: February 15, 2005, 18:22:09 ; Search time 714 Seconds
(without alignments)
3065.925 Million cell updates/sec

Title: US-08-876-132-3
Perfect score: 1947
Sequence: 1 NFLIHWKFEKIRQKKLKWKKY.........SQELLRLCSFVXRWYRCNWC 371

Scoring table: BLOSUM62.
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5378673 seqs, 2950229984 residues

Total number of hits satisfying chosen parameters: 10757346

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool_h/US08876132/runat_14022005_125917_22894/app_query.fasta_1.519
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US08876132 @CGN 1_1_480 @runat_14022005_125917_22894
-NCPU=6 -ICPU=3 -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	139.5	7.2	6497	18	US-10-719-993-248	Sequence 248, App
2	122.5	6.1	4728	18	US-10-437-963-20765	Sequence 20765, A
3	119	6.1	2726	17	US-10-282-122A-21447	Sequence 21447, A
4 c	119	6.0	1762	18	US-10-363-345A-7529	Sequence 7529, Ap
5	117.5	6.0	3471	17	US-10-369-493-23730	Sequence 23730, A
6 c	117.5	6.0	5418	17	US-10-282-122A-27329	Sequence 27329, A
7	117	6.0	580073	16	US-10-205-220-1	Sequence 1, Appli
8	117	6.0	2373	18	US-10-793-639-602	Sequence 602, App
9 c	116.5	6.0	4301	9	US-09-917-800A-1666	Sequence 1666, Ap
10	116.5	6.0	3027	15	US-10-205-194-126	Sequence 126, App
11	116	6.0	6334	15	US-10-282-122A-34957	Sequence 34957, A
12	116	6.0	2796	16	US-10-311-455-1186	Sequence 1186, Ap
13 c	116	6.0	4840	15	US-10-032-58S-6600	Sequence 6600, Ap
14	115	5.9	4840	18	US-10-437-963-72826	Sequence 72826, A
15	114.5	5.9	3534	17	US-10-282-122A-16250	Sequence 16250, A
16 c	114	5.9	6232	17	US-10-377-133-62	Sequence 62, Appl
17	114	5.9	6978	18	US-10-663-433-1	Sequence 1, Appl
18	114	5.9	7433	18	US-10-723-860-1449	Sequence 1449, Ap
19	114	5.9	1828	9	US-09-764-864-190	Sequence 190, App
20	113.5	5.8	2439	18	US-10-115-635-56	Sequence 56, Appli
21	113.5	5.8	3312	9	US-09-797-385-3	Sequence 3, Appli
22	113.5	5.8	3384	9	US-09-797-385-5	Sequence 5, Appli
23	113.5	5.8	5040	17	US-10-369-493-45767	Sequence 45767, A
24 c	113.5	5.8	6272	17	US-10-052-648A-19	Sequence 19, Appl
25	113.5	5.8	2256646	18	US-10-470-565-1	Sequence 1, Appli
26	113	5.8	6767	9	US-09-221-714A-330	Sequence 330, App
27 c	113	5.8	6773	9	US-09-864-864-336	Sequence 336, App
28	113	5.8	9399	17	US-10-396-122-93	Sequence 93, Appl
29	113	5.8	684707	17	US-10-398-221-9	Sequence 9, Appli
30	113	5.8	3011208	17	US-10-398-221-2058	Sequence 2058, Ap
31 c	113	5.8	2405	17	US-09-748-875-70	Sequence 70, Appl
32	112.5	5.8	2405	10	US-09-298-523B-70	Sequence 70, Appl
33	112.5	5.8	3018	17	US-10-369-493-24748	Sequence 24748, A
34	112.5	5.8	3025	17	US-09-960-706-972	Sequence 972, App
35	112.5	5.8	3025	10	US-09-873-319-634	Sequence 634, App
36	112.5	5.8	3025	14	US-10-171-311-47	Sequence 47, Appl
37	112.5	5.8	3025	18	US-10-788-792-39	Sequence 39, Appli
38	112.5	5.8	3294	9	US-09-797-385-7	Sequence 7, Appli
39	112.5	5.8	5601	17	US-10-424-599-133525	Sequence 133525,
40	112.5	5.8	885	17	US-10-369-493-37144	Sequence 37144, A
41	111.5	5.7	1635	18	US-10-437-963-40323	Sequence 40323, A
42	111.5	5.7	2278	18	US-10-425-115-170903	Sequence 170903, A
43	111.5	5.7	3423	17	US-10-369-493-41881	Sequence 41881, A
44	111.5	5.7	4200	9	US-09-797-385-1	Sequence 1, Appli

RESULT 1
US-10-719-993-248

; Sequence 248, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 248
; LENGTH: 6497
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-248

ALIGNMENTS
US-10-719-993-248

Alignment Scores:

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 15, 2005, 18:08:35 ; Search time 229 Seconds
               (without alignments)
               2650.911 Million cell updates/sec

Title: US-08-876-132-3
Perfect score: 1947
Sequence: 1 NFLHWKPEKIRQKKLKWKKY.........SQELLRLCSFVXRWYRCNWC 371

Searched: 1202784 seqs; 81818359 residues
Total number of hits satisfying chosen parameters: 2405568
               Listing first 45 summaries

Scoring table: BLOSUM62
               Xgapop 10.0 ; Xgapext 0.5
               Ygapop 10.0 ; Ygapext 0.5
               Fgapop 6.0 ; Fgapext 7.0
               Delop   6.0 ; Delext   7.0

Post-processing: Minimum Match 0%
                 Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US08876132/runat_14022005_125915_22812/app_query.fasta_1.519
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-USER=US08876132 @CGN_1_1_69 @runat_14022005 125915 22812 -NCPU=6 -ICPU=3
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000000
-NO_MMAP -LARGEQUERY_NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	130.5	6.7	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	2	130.5	6.7	1664976	4	US-09-692-570-1	Sequence 1, Appli
c	3	127.5	6.5	1449	4	US-09-248-796A-4693	Sequence 4693, Ap
	4	127.5	6.5	43576	4	US-09-676-519-19	Sequence 19, Appl
c	5	117	6.0	580073	4	US-08-545-528D-1	Sequence 1, Appli
	6	116.5	6.0	2373	3	US-09-614-221A-602	Sequence 602, App
	7	116	6.0	3033	3	US-09-134-001C-2341	Sequence 2341, Ap
	8	114	5.9	3881	4	US-09-949-016-3905	Sequence 3905, Ap
	9	114	5.9	3893	4	US-08-949-016-908	Sequence 908, App
	10	113.5	5.8	3312	4	US-08-923-992A-3	Sequence 3, Appli
	11	113.5	5.8	3384	3	US-08-923-992A-5	Sequence 5, Appli
	12	113	5.8	6773	3	US-09-166-350-27	Sequence 27, Appl
	13	112.5	5.8	3025	4	US-09-917-254-19	Sequence 19, Appl
	14	112.5	5.8	3294	4	US-08-923-992A-7	Sequence 7, Appli
	15	112.5	5.8	3763	1	US-07-792-865D-1	Sequence 1, Appli
	16	112.5	5.8	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	17	112.5	5.8	1664976	4	US-09-692-570-1	Sequence 1, Appli
	18	111.5	5.7	2617	1	US-08-430-024-1	Sequence 1, Appli
	19	111.5	5.7	2617	1	US-08-782-009-1	Sequence 1, Appli
	20	111.5	5.7	2617	3	US-09-017-302-1	Sequence 1, Appli
	21	111.5	5.7	3730	1	US-08-242-932-8	Sequence 8, Appli
	22	111.5	5.7	3730	1	US-08-714-481-8	Sequence 8, Appli
	23	111.5	5.7	3730	5	PCT-US95-06111-8	Sequence 8, Appli
	24	111.5	5.7	4200	1	US-08-242-932-1	Sequence 1, Appli
	25	111.5	5.7	4200	1	US-08-714-481-1	Sequence 1, Appli
	26	111.5	5.7	4200	5	PCT-US95-06111-1	Sequence 1, Appli
	27	111	5.7	4200	5	PCT-US95-06111-1	Sequence 1, Appli
	28	111	5.7	1497	3	US-08-946-026-58	Sequence 58, Appl
	29	111	5.7	9391	4	US-09-562-702A-11	Sequence 11, Appl
	30	111	5.7	9511	4	US-09-562-702A-9	Sequence 9, Appli
	31	110.5	5.7	5192	4	US-09-949-016-5831	Sequence 5831, Ap
	32	110	5.6	8257	4	US-00-595-684B-30	Sequence 30, Appl
	33	110	5.6	8503	4	US-09-620-312D-130	Sequence 130, Appl
	34	109	5.6	1176	4	US-09-107-532A-859	Sequence 859, App
	35	109	5.6	5053	2	US-09-685-576-2	Sequence 2, Appli
	36	108.5	5.6	1473	6	US-09-710-279-871	Sequence 871, App
	37	108.5	5.6	1536	4	US-09-248-796A-2277	Sequence 2277, App
	38	108.5	5.6	2169	4	US-09-248-796A-6510	Sequence 6510, Ap
c	39	108.5	5.6	3318	4	US-09-710-279-3869	Sequence 3869, Ap
c	40	108.5	5.6	3492	3	US-08-923-992A-9	Sequence 9, Appli
c	41	108.5	5.6	3607	4	US-09-710-279-4297	Sequence 4297, Ap
	42	108.5	5.6	3926	4	US-09-710-279-4015	Sequence 4015, Ap
	43	108	5.5	1266	4	US-09-248-796A-5218	Sequence 5218, Ap
c	44	107	5.5	789	4	US-09-248-796A-4956	Sequence 4956, Ap
	45	107	5.5	9626	4	US-09-150-867-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
Patent No. 6503729

GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
TITLE OF INVENTION: jannaschii
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 15, 2005, 13:11:01 ; Search time 680 Seconds
(without alignments)
3229.741 Million cell updates/sec

Title: US-08-876-132-3
Perfect score: 1947
Sequence: 1 NFLHWKPEKIRQKKLKWKKY...........SQELLRLCSFVXRWYRCNWC 371

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 295987O667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
N_Geneseq_16Dec04:*
-MODEL=frame+_p2n.model -DEV=-xlh
-Q=/cgn2_1/USPTO_spool_h/US08876132/runat_14022005_125913_22754/app_query.fasta_1.519
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000000
-USER=US08876132 @CCN_1_1_470 @runat_14022005_125913_22754_app
-NO_MMAP -LARGEQUERY_NEG_SCORES=0 -WAIT -DSPBLOCK100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	1943	99.8	1660	2	AAV74275	Pantoea c
c 2	148.5	7.6	35515	2	AAX20252	Borrelia
c 3	141.5	7.3	3300	13	ADR08302	Full leng
c 4	139.5	7.2	3138	4	AAH46300	Pseudoalt
c 5	130.5	6.7	110000	2	AAV21209_11	Continuation (12 o
c 6	129.5	6.7	26776	2	AAX20254	Borrelia
c 7	129	6.7	7078	4	ABL03681	Drosophil
c 8	129	6.6	13085	4	ABL03680	Drosophil
c 9	127.5	6.6	43576	3	AAA68609	Bacteriop
c 10	127	6.5	110000	2	AAX20248_03	Continuation (4 of
c 11	124	6.4	4663	2	AAX20273	Borrelia
c 12	122	6.3	2562	4	ABL09845	Drosophil
c 13	121.5	6.2	1449	3	AAA70251	Plasmodiu
c 14	120.5	6.2	5994	8	AAA70222	Plasmodiu
c 15	119	6.1	2726	8	ACA33577	Prokaryot
c 16	118.5	6.1	2492	2	AAQ14755	FUS2 gene
c 17	117.5	6.0	1762	6	ABQ20939	Oligonucl
c 18	117.5	6.0	1762	6	ABQ20938	Oligonucl
c 19	117.5	6.0	3471	13	ADS45300	Bacterial
c 20	117	6.0	2943	8	AAA70229	Prokaryot
c 21	117	6.0	5418	8	ACA39459	Prokaryot
c 22	117	6.0	6016	4	AAH57390	Human ske
c 23	117	6.0	110000	2	AAT58840_2	Continuation (3 of
c 24	117	6.0	3669	8	ADA71351	Rice gene
c 25	116.5	6.0	4301	6	ABK63759	Rat seque
c 26	116.5	6.0	4301	10	ADB58345	Toxicity
c 27	116.5	6.0	4301	10	ADB52908	Primary r
c 28	116.5	6.0	4301	10	ADB85245	Rat astro
c 29	116.5	6.0	7620	2	AAA70132	Plasmodiu
c 30	116.5	6.0	24789	4	ABL28640	Drosophil
c 31	116	6.0	3027	8	ACA47087	Prokaryot
c 32	116	6.0	3033	5	ABN92878	Staphyloc
c 33	116	6.0	3033	13	ADS01110	Staphyloc
c 34	116	6.0	3951	4	ABL10543	Drosophil
c 35	116	6.0	4702	4	ABL09844	Drosophil
c 36	116	6.0	6334	6	ABL33213	Human imm
c 37	116	6.0	89378	12	ADN46845_20	Continuation (21 o
c 38	116	6.0	89378	12	ADN46123_20	Continuation (21 o
c 39	116	6.0	89378	12	ADN46464_20	Continuation (21 o
c 40	116	6.0	110000	12	ADN46845_19	Continuation (20 o
c 41	116	6.0	110000	12	ADN47591_00	Thermococ
c 42	116	6.0	3033	13	ADN46123_19	Continuation (20 o
c 43	116	6.0	3951	4	ADN47209_00	Thermococ
c 44	116	6.0	110000	12	ADN46464_19	Continuation (20 o
c 45	116	6.0	110000	12	ADN47960_00	Thermococ

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

Aax20254 Borrelia
Ab103681 Drosophil
Ab103680 Drosophil
Aaa68609 Bacteriop
Continuation (4 of
Aax20273 Borrelia
Ab109845 Drosophil
Aaa70251 Plasmodiu
Aaa70222 Plasmodiu
Aca33577 Prokaryot
Aaq14755 FUS2 gene
Abq20939 Oligonucl
Abq20938 Oligonucl
Ads45300 Bacterial
Aaa70229 Prokaryot
Aca39459 Prokaryot
Aah57390 Human ske
Continuation (3 of
Ada71351 Rice gene
Abk63759 Rat seque
Adb58345 Toxicity
Adb52908 Primary r
Adb85245 Rat astro
Aaa70132 Plasmodiu
Ab128640 Drosophil
Aca47087 Prokaryot
Abn92878 Staphyloc
Ads01110 Staphyloc
Ab110543 Drosophil
Ab109844 Drosophil
Ab133213 Human imm
Continuation (21 o
Continuation (21 o
Continuation (21 o
Continuation (20 o
Continuation (20 o
Continuation (20 o
Adn47209_00 Thermococ
Continuation (20 o
Adn47960_00 Thermococ

RESULT 1
AAV74275

ID AAV74275 standard; DNA; 1660 BP.
XX
AC AAV74275;
XX
DT 16-JUN-1999 (first entry)
XX
DE Pantoea citrea cryptic plasmid pS.
XX
KW Cryptic plasmid; pS; higher temperature; growth; elimination;
KW mobilization; ds.
XX
OS Pantoea citrea.
XX
FH Key Location/Qualifiers
FT CDS 480..1595
FT /*tag= a
FT /note= "largest open reading frame"
XX
PN WO9859054-A1.
XX
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98WO-US0129945.
XX
PR 23-JUN-1997; 97US-00876132.
XX
PA (GEMV) GENENCOR INT INC.

```
OM protein - nucleic search, using frame_plus_p2n model

                                          GenCore version 5.1.6
                                          Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:           February 15, 2005, 13:21:06 ; Search time 5433 Seconds
                                                (without alignments)
                                                3308.832 Million cell updates/sec

Title:            US-08-876-132-3
Perfect score:    1947
Sequence:         1 NFLHWKFEKIRQKKLKWKKY..........SQELLRLCSFVXRWYRCNWC 371

Scoring table:    BLOSUM62
                  Xgapop 10.0 , Xgapext  0.5
                  Ygapop 10.0 , Ygapext  0.5
                  Fgapop  6.0 , Fgapext  7.0
                  Delop   6.0 , Delext   7.0

Searched:         4708233 seqs, 2422607955 residues

Total number of hits satisfying chosen parameters:   9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model
-Q=/cgn2_1/USPTO_spool_h/US08876132//runat_14022005_125914_22772/app_query.fasta_1.519
-DB=GenEmbl -QFMT=fastp -SUFFIX=rge -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US08876132 @CGN_1_1_3731 @runat_14022005_125914_22772 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :        GenEmbl:

                  1:  gb_ba:*
                  2:  gb_htg:*
                  3:  gb_in:*
                  4:  gb_om:*
                  5:  gb_ov:*
                  6:  gb_pat:*
                  7:  gb_ph:*
                  8:  gb_pl:*
                  9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.


                                    SUMMARIES

                             %
Result        Query  Match  Length  DB  ID          Description
No.    Score
-----  -----  -----  -----  ------  --  --          -----------
   1  1643.5   84.4   3661  3661   1   AF128889    AF128889 Pantoea c
   2   170.5    8.8  113880  3   1   PFMAL3P4    PFMAL3P4 Plasmodiu
   3   164.5    8.4  30299   1   1   AE001577    AE001577 Borrelia
   4   155.5    8.0   3871   1   1   BBRGBCDE    X87202 Borrelia bu
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2005, 17:11:45 ; Search time 4043 Seconds
                                     (without alignments)
                                     3492.912 Million cell updates/sec

Title: US-08-876-132-3
Perfect score: 1947
Sequence: 1 NFLHWKFEKIRQKKLKWKKY..........SQELLRLCSFVXRWYRCNWC 371

Scoring table: BLOSUM62
               Xgapop   10.0 , Xgapext   0.5
               Ygapop   10.0 , Ygapext   0.5
               Fgapop   6.0 , Fgapext   7.0
               Delop   6.0 , Delext   7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US0876132//runat_14022005_125914_22795/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000000
-USER=US0876132_@CGN_1_3437_@runat_14022005_125914_22795 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
           1: gb_est1:*
           2: gb_est2:*
           3: gb_htc:*
           4: gb_est3:*
           5: gb_est4:*
           6: gb_est5:*
           7: gb_est6:*
           8: gb_ss1:*
           9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

				%			
Result No.	Score	Query Match	Length	DB	ID	Description	
1	132	6.8	711	9	CE673882	CE673882 tigr-gss-
2	131.5	6.8	747	9	CE624742	CE624742 tigr-gss-
3	130.5	6.7	721	9	CE561141	CE561141 tigr-gss-
4	129	6.6	678	9	CE558403	CE558403 tigr-gss-
5	127	6.5	604	9	CE350303	CE350303 tigr-gss-
6	126.5	6.5	723	9	CE608358	CE608358 tigr-gss-
7	125.5	6.4	634	9	CE216131	CE216131 tigr-gss-
8	125.5	6.4	650	9	CE165694	CE165694 tigr-gss-
9	125.5	6.4	658	9	CE589588	CE589588 tigr-gss-
10	125.5	6.4	2740	3	AK029960	AK029960 Mus muscu
c 11	124	6.4	571	9	CE786366	CE786366 tigr-gss-
12	124	6.4	696	9	CE847184	CE847184 tigr-gss-
13	123.5	6.4	672	9	CE200159	CE200159 tigr-gss-
14	123.5	6.3	786	9	CE677943	CE677943 tigr-gss-
15	123	6.3	583	9	CE524900	CE524900 tigr-gss-
16	123	6.3	598	9	CE591938	CE591938 tigr-gss-
17	123	6.3	621	9	CE718400	CE718400 tigr-gss-
18	123	6.3	690	9	CE226283	CE226283 tigr-gss-
19	123	6.3	695	9	CE568932	CE568932 tigr-gss-
20	123	6.3	698	9	CE616081	CE616081 tigr-gss-
21	123	6.3	703	9	CE575059	CE575059 tigr-gss-
22	123	6.3	708	9	CE012114	CE012114 tigr-gss-
23	123	6.3	712	9	CE801588	CE801588 tigr-gss-
c 24	123	6.3	718	9	CE010297	CE010297 tigr-gss-
25	123	6.3	755	9	CE038597	CE038597 tigr-gss-
c 26	122.5	6.3	646	9	CE537601	CE537601 tigr-gss-
c 27	122.5	6.3	667	9	CE378783	CE378783 tigr-gss-
28	122.5	6.3	668	9	CE147044	CE147044 tigr-gss-
29	122.5	6.3	673	9	CE690020	CE690020 tigr-gss-
30	122.5	6.3	684	9	CE194890	CE194890 tigr-gss-
31	122.5	6.3	720	9	CE759049	CE759049 tigr-gss-
32	122.5	6.3	734	9	CE728920	CE728920 tigr-gss-
33	122.5	6.3	738	9	CE438325	CE438325 tigr-gss-
34	122.5	6.3	742	9	CE800878	CE800878 tigr-gss-
35	122.5	6.3	742	9	CE439065	CE439065 tigr-gss-
36	122.5	6.3	749	9	CE491787	CE491787 tigr-gss-
37	122	6.3	644	9	CE083270	CE083270 tigr-gss-
c 38	122	6.3	696	9	CE196009	CE196009 tigr-gss-
39	122	6.3	713	9	CE822200	CE822200 tigr-gss-
40	122	6.3	740	9	CE635780	CE635780 tigr-gss-
41	122	6.3	773	9	CE103096	CE103096 tigr-gss-
42	122	6.3	818	9	CE035177	CE035177 tigr-gss-
c 43	121.5	6.2	735	9	CE778946	CE778946 tigr-gss-
c 44	121	6.2	688	9	CE469175	CE469175 tigr-gss-
c 45	121	6.2	692	9	CE469175	CE469175 tigr-gss-

ALIGNMENTS

RESULT 1
CE673882
LOCUS   CE673882   711 bp   DNA   linear   GSS 29-SEP-2003
DEFINITION tigr-gss-dog-1700032944777791 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE673882
VERSION   CE673882.1   GI:36992882
KEYWORDS   GSS.
SOURCE   Canis familiaris (dog)
  ORGANISM Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1 (bases 1 to 711)
  AUTHORS   Kirkness,E.F., Bazna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE   The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
  MEDLINE   22875432
   PUBMED   14512627
COMMENT   Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirkness@tigr.org
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..711
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 15, 2005, 09:55:11 ; Search time 1290.32 seconds
(without alignments)
8446.087 Million cell updates/sec

Title: US-08-876-132-2
Perfect score: 1847
Sequence: 1 AGATCTCAACCAGTTTAAAA.........AAACAAGCAAAAAGAGATCT 1847

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5378673 seqs, 295029984 residues

Total number of hits satisfying chosen parameters: 10757346

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	length	DB ID	ID	Description
c 1	130	7.0	7597	15	US-10-311-455-986	Sequence 986, App
c 2	123.2	6.7	6286	17	US-10-221-714A-313	Sequence 313, App
c 3	114.2	6.2	6109	15	US-10-311-455-299	Sequence 299, App
c 4	114.2	6.2	6109	17	US-10-221-613-33	Sequence 33, Appl
c 5	113.8	6.2	7814	17	US-10-221-714A-252	Sequence 252, App
c 6	111.6	6.0	12237	15	US-10-311-455-2331	Sequence 2331, App
c 7	101.2	5.5	363778	16	US-10-312-841-1	Sequence 1, Appli
c 8	97.8	5.3	13321	17	US-10-221-714A-143	Sequence 143, App
c 9	97.8	5.3	163382	18	US-10-367-094-179	Sequence 179, App
c 10	97.4	5.3	158001	17	US-10-211-179-11	GENERAL INFORMATI
c 11	97	5.3	5413	17	US-10-221-714A-418	Sequence 418, App
c 12	96.2	5.2	6419	15	US-10-311-455-240	Sequence 240, App
c 13	94	5.1	8056	16	US-10-473-126-386	Sequence 386, App
c 14	93	5.0	46951	18	US-10-091-281-2	Sequence 2, Appli
c 15	91.4	4.9	6154	17	US-10-221-613-69	Sequence 69, Appl
c 16	91.2	4.9	6292	17	US-10-221-714A-461	Sequence 461, App
c 17	90.2	4.9	7442	17	US-10-221-714A-409	Sequence 409, App
c 18	90.2	4.9	8056	18	US-10-473-126-386	Sequence 386, App
c 19	89.6	4.9	6048	18	US-10-433-793-32	Sequence 32, Appl
c 20	89.6	4.9	6106	18	US-10-311-455-1445	Sequence 1445, Ap
c 21	89.6	4.9	6106	17	US-10-257-166-113	Sequence 113, App
c 22	89.6	4.9	6106	17	US-10-221-714A-151	Sequence 151, App
c 23	88.4	4.8	15548	17	US-10-311-455-2128	Sequence 2128, Appli
c 24	88.4	4.8	3673778	16	US-10-312-841-2	Sequence 2, Appli
c 25	88.2	4.8	1300	18	US-10-668-749A-1	Sequence 1, Appli
c 26	86.6	4.7	1243	18	US-10-473-126-240	Sequence 240, App
c 27	86.2	4.7	9760	17	US-10-221-613-113	Sequence 113, App
c 28	86.2	4.7	8056	18	US-10-311-455-1722717	Sequence 172717
c 29	86	4.7	6175	15	US-10-311-455-1280	Sequence 1280, Ap
c 30	85.8	4.6	3683	18	US-10-473-126-193	Sequence 193, App
c 31	85.8	4.6	3683	18	US-10-473-126-339	Sequence 339, App
c 32	85.4	4.6	8056	18	US-10-473-126-240	Sequence 240, App
c 33	85.4	4.6	7498	15	US-10-311-455-230	Sequence 230, App
c 34	84.6	4.6	10048	17	US-10-311-455-211	Sequence 211, App
c 35	84	4.5	6831	15	US-10-221-613-211	Sequence 211, App
c 36	82.8	4.5	921	18	US-10-425-115-38710	Sequence 38710, A
c 37	82	4.4	49979	18	US-10-741-601-5746	Sequence 5746, Ap
c 38	82	4.4	49979	19	US-10-239-676-90	Sequence 90, Appl
c 39	81.6	4.4	5689	14	US-10-240-453-100	Sequence 100, App
c 40	81.6	4.4	5689	15	US-10-221-714A-148	Sequence 148, App
c 41	81.6	4.4	5689	17	US-10-367-094-179	Sequence 179, App
c 42	81.6	4.4	163382	18	US-10-425-115-16756	Sequence 16756, A
c 43	81.4	4.4	1081	18	US-10-240-453-302	Sequence 302, App
c 44	81	4.4	8961	15	US-10-311-455-330	Sequence 330, App
c 45	81	4.4	9642	15	US-10-311-455-330	Sequence 418, App

ALIGNMENTS

RESULT 1
US-10-311-455-986/c
Sequence 986, Application US/10311455
Publication No. US20030143606A1
US-311-455-986
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter m
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 986
; LENGTH: 7597
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-986

Query Match   7.0%;   Score 130;   DB 15;   Length 7597;
Best Local Similarity   47.7%;   Pred. No. 1.6e-12;
Matches 447;   Conservative 0;   Mismatches 485;   Indels 6;   Gaps 2;

Qy   912 ATACCAGRAGCAAAACAAAAAAAATAAAACAAAGAAAAAAATTTTCGAGCGAAAAAAATATTT 971

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:　February 15, 2005, 03:19:09 ; Search time 313.89 seconds
　　　　　(without alignments)
　　　　　9628.226 Million cell updates/sec

Title:　　　　US-08-876-132-2
Perfect score:　1847
Sequence:　　1 AGATCTCAACCAGTTTAAAA.................AAACAAGCAAAAAGAGATCT 1847
Scoring table:　IDENTITY NUC
　　　　　Gapop 10.0 , Gapext 1.0
Searched:　1202784 seqs, 81138359 residues

Total number of hits satisfying chosen parameters:　2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
　　　　　Maximum Match 100%
　　　　　Listing first 45 summaries

Database :　Issued Patents NA:*
　　1:/cgn2_6/ptodata/1/ina/5A_COMB.seq:*
　　2:/cgn2_6/ptodata/1/ina/5B_COMB.seq:*
　　3:/cgn2_6/ptodata/1/ina/6A_COMB.seq:*
　　4:/cgn2_6/ptodata/1/ina/6B_COMB.seq:*
　　5:/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
　　6:/cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	97.8	5.3	147382	4	US-09-949-016-14624	Sequence 14624, A
2	88.6	4.8	187169	4	US-09-949-016-12776	Sequence 12776, A
3	88.6	4.8	191569	4	US-09-949-016-15940	Sequence 15940, A
4	84.8	4.6	134987	4	US-09-949-016-15348	Sequence 15348, A
5	84.8	4.6	134987	4	US-09-949-016-15349	Sequence 15349, A
6	84.8	4.6	134987	4	US-09-949-016-15350	Sequence 15350, A
7	84.8	4.6	134987	4	US-09-949-016-15507	Sequence 15507, A
8	84.8	4.6	134987	4	US-09-949-016-15508	Sequence 15508, A
9	84.8	4.6	134987	4	US-09-949-016-15509	Sequence 15509, A
10	82.6	4.5	29717	4	US-09-949-016-16284	Sequence 16284, A
11	82.6	4.5	60376	4	US-09-949-016-12423	Sequence 12423, A
12	82	4.4	18773	4	US-09-949-016-14164	Sequence 14164, A
13	81.6	4.4	147382	4	US-09-949-016-14624	Sequence 14624, A
14	80	4.3	205044	4	US-09-949-016-15851	Sequence 15851, A
15	80	4.3	205044	4	US-09-949-016-15852	Sequence 15852, A
16	80	4.3	205044	4	US-09-949-016-15853	Sequence 15853, A
17	80	4.3	223471	4	US-09-949-016-12387	Sequence 12387, A
18	80	4.3	223471	4	US-09-949-016-12724	Sequence 12724, A
19	80	4.3	223471	4	US-09-949-016-12725	Sequence 12725, A
20	80	4.2	18773	4	US-09-949-016-14164	Sequence 14164, A
21	77	4.2	612	4	US-09-902-540-1357	Sequence 1357, Ap
22	75.4	4.1	1039	4	US-09-902-540-1280	Sequence 1280, Ap
23	75	4.1	601	4	US-09-949-016-30531	Sequence 30531, A
24	75	4.1	601	4	US-09-949-016-37150	Sequence 37150, A
25	75	4.1	601	4	US-09-949-016-37164	Sequence 37164, A
26	75	4.1	601	4	US-09-949-016-145868	Sequence 145868, A
27	75	4.1	601	4	US-09-949-016-146136	Sequence 146136,
28	74.8	4.1	601	4	US-09-949-016-146404	Sequence 146404,
29	74.8	4.0	95255	4	US-09-949-016-17067	Sequence 17067, A
30	74.6	4.0	19124	2	US-08-487-826B-13	Sequence 13, Appl
31	74.2	4.0	251672	4	US-09-949-016-17296	Sequence 17296, A
32	74.2	4.0	251682	4	US-09-949-016-11973	Sequence 11973, A
33	73.4	4.0	134987	4	US-09-949-016-15348	Sequence 15348, A
34	73.4	4.0	134987	4	US-09-949-016-15349	Sequence 15349, A
35	73.4	4.0	134987	4	US-09-949-016-15350	Sequence 15350, A
36	73.4	4.0	134987	4	US-09-949-016-15507	Sequence 15507, A
37	73.4	4.0	134987	4	US-09-949-016-15508	Sequence 15508, A
38	73.4	3.9	263694	4	US-09-949-016-12386	Sequence 12386, A
39	72.6	3.9	263693	4	US-09-949-016-16915	Sequence 16915, A
40	72.6	3.9	12313	4	US-09-949-016-13248	Sequence 13248, A
41	72	3.9	12313	4	US-09-949-016-12378	Sequence 12378, A
42	71.2	3.9	119153	4	US-09-949-016-17145	Sequence 17145, A
43	71	3.8	30820	4	US-09-949-016-12776	Sequence 12776, A
44	70	3.8	187169	4	US-09-949-016-12776	Sequence 12776, A
45	70	3.8	191569	4	US-09-949-016-15940	Sequence 15940, A

ALIGNMENTS

RESULT 1
US-09-949-016-14624/c
; Sequence 14624, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14624
; LENGTH: 147382
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(147382)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14624

Query Match　　　5.3%;　Score 97.8;　DB 4;　Length 147382;
Best Local Similarity　48.4%;　Pred. No. 1.1e-10;
Matches 328; Conservative 0; Mismatches 347; Indels 2; Gaps 2;

Qy　1005 TTTTGCCATATTTAAAACCTGACTATCTTTATAAGTTAAGTTAATAATCCGTTAGATTAT 1064
　　　　　 ||||||||| || |||||||||||||||| || |||||||| || ||||
Db 137331 TTTCACGAAATTTAAACTCTCTCTCTCTAAAATAATATTTATAAATATATATATAAATAA 137272

Qy　1065 AAAGTATGTTAAAAACCAGTAAAAACAATAACTTTATAATATATTTAAATTCTGAATTATATATTTG 1124
　　　　　 |||||||| |||||||| ||||||||||||||||||| ||||| |||||||
Db 137271 ATATATAAATATATATAAATATAAAATATAAAATAAATAAATATAAATATAAAATATATAAA 137212

Qy　1125 ACAGTGATTATTTAATTATATTATTAAGAGATATATCTATTAGCTTAAATATTAAAAAAAG 1184
　　　　　 ||||||||||||||||||| || |||||||| |||||||||||||||
Db 137211 ATATATAAATATAAATAAATAAATATAAAAATATAAATAAATATAAAATAAATATAAATAT 137152

Qy　1185 AGGTAAATATATGGATTGTGTATTTAAAAAGCATTAGAAAAATGAAATAGAACATTATAA 1244
　　　　　 ||||| || ||||||| || ||| |||||| || |||||||
Db 137151 ATAAATATGTTATAAATACATATATAAATAAATAATAATATATATAAATATATATATCAT 137092

Qy　1245 AAAAGACGGGTGATATCAAAATCTTTCTTACAAATACTTGCATTACTTTGATATAGATAAAGC 1304

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 15, 2005, 00:27:38 ; Search time 1003.82 seconds
(without alignments)
10892.200 Million cell updates/sec

Title: US-08-876-132-2
Perfect score: 1847
Sequence: 1 AGATCTCAACCAGTTTAAAA...........AAACAAGCAAAAAGAGATCT 1847

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :
       1: N_Geneseq_16Dec04:*
       2: geneseqn1980s:*
       3: geneseqn1990s:*
       4: geneseqn2000s:*
       5: geneseqn2001as:*
       6: geneseqn2001bs:*
       7: geneseqn2002as:*
       8: geneseqn2002bs:*
       9: geneseqn2003as:*
      10: geneseqn2003bs:*
      11: geneseqn2003cs:*
      12: geneseqn2003ds:*
      13: geneseqn2004as:*
          geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	1847	100.0	1847	2	AAV74276	Pantoea c
c 2	130	7.0	7597	6	ABL33013	Human imm
c 3	123.2	6.7	6286	6	AAS46591	Tumour su
c 4	114.2	6.2	6109	6	ABL32326	Human imm
c 5	114.2	6.2	6109	6	AAS61077	Human gen
c 6	113.8	6.2	7814	4	AAS46530	Tumour su
c 7	111.6	6.0	12237	6	ABL34358	Human imm
c 8	99.8	5.4	8222	8	ACF62794	Colon can
c 9	99.8	5.4	11222	10	ADB54190	Pretreate
c 10	99.8	5.4	8222	8	ACF62816	Colon can
c 11	98.2	5.3	11222	10	ADB54318	Pretreate
c 12	98.2	5.3	8222	8	AAK73165	Human imm
c 13	98.2	5.3	20420	6	ABK69933	Human imm
c 14	97.8	5.3	13321	13	AAS46421	Tumour su
c 15	97.8	5.3	163382	12	ABD32659	Human can
c 16	97.4	5.3	158001	12	ADL17884	Human pho
c 17	97	5.3	5413	4	AAS46694	Tumour su
c 18	96.4	5.2	19965	4	AAK73166	Human imm
c 19	96.4	5.2	19965	4	ABK69932	Human sec
c 20	96.2	5.2	6419	6	ABL32267	Human imm
c 21	94.2	5.1	115218	8	ACA64845	Human HNR
c 22	94	5.1	8056	8	AB210246	Haematopo
c 23	93	5.0	46951	10	ADE13891	Human opt
c 24	91.4	4.9	8056	8	ABK31200	Signal tr
c 25	91.4	4.9	6154	6	ABL70167	Chemical
c 26	91.2	4.9	6154	6	AAS61112	Human gen
c 27	91.2	4.9	6292	4	AAS46735	Tumour su
c 28	91.2	4.9	110000	10	ADH10017_0	Human chr
c 29	90.2	4.9	7442	4	AAS46686	Tumour su
c 30	90.2	4.9	8056	8	AB210246	Haematopo
c 31	89.6	4.9	6048	6	ABQ67002	Human ang
c 32	89.6	4.9	6106	4	AAS46429	Tumour su
c 33	89.6	4.9	6106	6	ABK40031	Human che
c 34	89.6	4.9	6106	6	ABL33472	Human imm
c 35	88.8	4.8	5286	6	ADS89278	Oligonucl
c 36	88.8	4.8	5286	6	ADS89552	Oligonucl
c 37	88.6	4.8	15548	6	ABL34155	Human imm
c 38	88.2	4.8	1000	12	ADQ62833	Homopoly-
c 39	88.2	4.8	1000	12	ADQ62832	Homopoly-
c 40	88.2	4.8	1300	12	ADP85917	Synthetic
c 41	87.4	4.7	110000	12	ABD32968_6	Continuation (7 of
c 42	86.6	4.7	8056	8	AB210100	Haematopo
c 43	86.2	4.7	9760	6	ABK31242	Signal tr
c 44	86.2	4.7	9760	6	ABL70197	Chemical
c 45	86.2	4.7	9760	6	AAS61155	Human gen

ALIGNMENTS

RESULT 1
AAV74276

ID   AAV74276 standard; DNA; 1847 BP.
XX
AC   AAV74276;
XX
DT   16-JUN-1999 (first entry)
XX
DE   Pantoea citrea cryptic plasmid pS.
XX
KW   Cryptic plasmid; pS; higher temperature; growth; elimination;
KW   mobilization; ds.
XX
OS   Pantoea citrea.
XX
PN   WO9859054-A1.
XX
PD   30-DEC-1998.
XX
PF   22-JUN-1998; 98WO-US012945.
XX
PR   23-JUN-1997; 97US-00876132.
XX
PA   (GEMV) GENENCOR INT INC.
XX
PI   Fowler T, Causey SC;
XX
DR   WPI; 1999-105625/09.
XX
PT   Method for preparing an improved Enterobacteriaceae strain - useful for
PT   improving bacterial fermentation strains of the family
PT   Enterobacteriaceae.
XX
PS   Claim 19; Fig 1D-F; 30pp; English.
XX
CC   The sequence is that of one half of the cryptic plasmid pS from Pantoea
CC   citrea. It can be used as part of a method of eliminating the cryptic
CC   plasmid from an Enterobacteriaceae strain which is useful for reducing
CC   the mobilization properties of plasmids residing within these strains.
CC   The elimination of the cryptic plasmid from the Enterobacteriaceae strain
CC   permits growth of the organism at a higher temperature, which decreases
CC   the production time for desired compounds in the carbohydrate pathway. It
CC   also has the commercial benefit of reducing both capital cost and

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 15, 2005, 00:33:38 ; Search time 7976.28 Seconds
(without alignments)
11220.366 Million cell updates/sec

Title: US-08-876-132-2
Perfect score: 1847
Sequence: 1 AGATCTCAACCAGTTTAAAA.........AAACAAGCAAAAAAGAGATCT 1847

Scoring table: IDENTITY.NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422607955 residues

Total number of hits satisfying chosen parameters: 9416466

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID		Description
1	1066.6	57.7	3661	1	AF128889	AF128889	Pantoea c
2	130	7.0	7597	6	AX345915	AX345915	Sequence
3	123.2	6.7	6286	6	AX251345	AX251345	Sequence
4	114.2	6.2	6109	6	AX251772	AX251772	Sequence
5	114.2	6.2	6109	6	AX345228	AX345228	Sequence
6	114.2	6.2	111861	9	AC069435	AC069435	Homo sapi
7	113.8	6.2	7814	6	AX251284	AX251284	Sequence
8	112.8	6.1	125632	2	CR753868	CR753868	Danio rer
9	112.6	6.1	171105	3	AC007532	AC007532	Drosophil
10	112.6	6.1	190801	3	AC009212	AC009212	Drosophil
11	112.6	6.1	295289	3	AE003603	AE003603	Drosophil
12	112.2	6.1	170627	2	AC125567	AC125567	Rattus no
13	111.6	6.0	12237	6	AX347260	AX347260	Sequence
14	111.6	6.0	174001	2	BX927081	BX927081	Danio rer
15	110.8	6.0	250022	3	AE014824	AE014824	Plasmodiu
16	110	6.0	67970	3	AL031746	AL031746	Plasmodiu
17	109.2	5.9	28150	3	AC137733	AC137733	Homo sapi
18	109	5.9	14867	3	AE001398	AE001398	Plasmodiu
19	108.8	5.9	90550	9	AL592166	AL592166	Human DNA
20	108.8	5.9	210548	2	CR751608	CR751608	Danio rer
21	107	5.8	125836	2	AC137629	AC137629	Homo sapi
22	106.8	5.8	175544	2	AC117342	AC117342	Rattus no
23	106.8	5.8	213056	9	AC103590	AC103590	Homo sapi
24	106.2	5.7	145650	2	CR762412	CR762412	Danio rer
25	106.2	5.7	148782	9	AP006213	AP006213	Homo sapi
26	105.2	5.7	161078	9	AP003500	AP003500	Homo sapi
27	105.2	5.7	161286	2	AC025120	AC025120	Homo sapi
28	105.2	5.7	180629	2	CR388025	CR388025	Danio rer
29	104.4	5.7	219074	2	CR769779	CR769779	Danio rer
30	104.4	5.7	164772	2	CR382291	CR382291	Danio rer
31	104.2	5.7	176704	2	CR628364	CR628364	Danio rer
32	104.2	5.6	162666	9	AC125796	AC125796	Homo sapi
33	104	5.6	165959	2	CR792429	CR792429	Danio rer
34	103.4	5.6	100925	9	AC012627	AC012627	Homo sapi
35	103.4	5.6	120481	9	AC099738	AC099738	Homo sapi
36	103.4	5.6	135121	9	AC069525	AC069525	Homo sapi
37	103.4	5.6	160667	2	AC151623	AC151623	Homo sapi
38	103.4	5.6	235532	9	AC008739	AC008739	Homo sapi
39	103.2	5.6	79018	9	AL627107	AL627107	Human DNA
40	103	5.6	171537	2	CR762390	CR762390	Danio rer
41	102.6	5.6	37330	9	AC133133	AC133133	Homo sapi
42	102.6	5.6	151032	9	AC105201	AC105201	Homo sapi
43	102.4	5.5	89807	2	CR536602	CR536602	Danio rer
44	102.4	5.5	262608	2	AC092362	AC092362	Danio rer
45	101.8	5.5	131682	9	AL672277	AL672277	Human DNA

ALIGNMENTS

RESULT 1
AF128889

```
LOCUS       AF128889                3661 bp    DNA     circular BCT 22-MAR-1999
DEFINITION  Pantoea citrea strain ATCC31623 plasmid pPZG500, complete plasmid
            sequence.
ACCESSION   AF128889
VERSION     AF128889.1  GI:4457229
KEYWORDS    .
SOURCE      Pantoea citrea
  ORGANISM  Pantoea citrea
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Pantoea.
REFERENCE   1  (bases 1 to 3661)
  AUTHORS   Bilic,M. and Delic,V.
  TITLE     Isolation and characterization of a cryptic plasmid from Erwinia
            citreus ATCC 31623
  JOURNAL   J. Appl. Microbiol. 83 (4), 485-492 (1997)
  MEDLINE   98012514
   PUBMED   9351229
REFERENCE   2  (bases 1 to 3661)
  AUTHORS   Bilic Nezic,M. and Delic,V.
  TITLE     Sequence Analysis and Functional Characterization of Replication
            Region in Plasmid pPZG500 from Pantoea citrea ATCC 31623
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 3661)
  AUTHORS   Bilic Nezic,M. and Delic,V.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-1999) Biotechnology, PLIVA d.d., Research
            Institute, Prilaz baruna Filipovica 25, Zagreb 10 000, Croatia
FEATURES             Location/Qualifiers
     source          1..3661
                     /organism="Pantoea citrea"
                     /mol_type="genomic DNA"
                     /strain="ATCC31623"
                     /db_xref="ATCC:31623"
                     /db_xref="taxon:53336"
                     /plasmid="pPZG500"
     misc_RNA        complement(join(3188..3661,1..177))
                     /note="RNAII"
     CDS             99..317
                     /note="Orf1"
                     /codon_start=1
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2005, 00:53:24 ; Search time 6187.74 Seconds
(without alignments)
11361.937 Million cell updates/sec

Title: US-08-876-132-2
Perfect score: 1847
Sequence: 1 AGATCTCAACCAGTTTAAAA.........AAACAAGCAAAAGAGATCT 1847

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34329544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	121.6	6.6	1348	9	CG749499	CG749499 P043-4-A0
2	114.8	6.2	1210	9	CG749728	CG749728 P044-1-C0
3	114.8	6.2	1805	9	CL080711	CL080711 CH216-159
4	114	6.2	1539	9	AG340947	AG340947 Mus muscu
5	112.8	6.1	1256	9	CL119201	CL119201 ISB1-76J1
6	112.2	6.1	1392	9	CG757503	CG757503 P052-4-C0
7	111.2	6.0	1626	7	CF238805	CF238805 AGENCOURT
8	108.2	5.9	1542	9	AG386981	AG386981 Mus muscu
9	107.8	5.8	1101	9	CNS00EVL	CNS00EVL Drosophil
10	107.8	5.8	1217	9	AL069706	AL069706 Drosophil
11	107.6	5.8	1599	9	CL062848	CL062848 CH216-98N
12	106.8	5.8	1042	9	CL083840	CL083840 ISB1-2H14
13	106.4	5.8	1594	9	CL461494	CL461494 SAIL_1148
14	106.2	5.7	1608	9	CL110653	CL110653 ISB1-53P2
15	106.2	5.7	1981	9	CL118721	CL118721 ISB1-72J8
16	106	5.7	1536	9	CL082000	CL082000 CH216-165
17	105.4	5.7	1355	9	CL078538	CL078538 CH216-151
18	104.8	5.7	1459	9	AG346348	AG346348 Mus muscu
19	104.6	5.7	1632	9	CL040025	CL040025 CH216-49A
20	104.2	5.6	651	9	CL082569	CL082569 CH216-167
21	104.2	5.6	659	9	CG807569	CG807569 1118085A0
22	104	5.6	639	9	CG732143	CG732143 1119146D0
23	104	5.6	1401	9	CG721246	CG721246 1119066B0
24	103.8	5.6	1243	9	CL016663	CL016663 CH216-1A2
25	103.6	5.6	1519	9	AG386893	AG386893 Mus muscu
26	103.2	5.6	1842	9	CL044158	CL044158 CH216-59P
27	103	5.6	1353	9	CG744812	CG744812 P037-3-B0
28	103	5.6	1784	9	CL081992	CL081992 CH216-165
29	102.8	5.6	1433	9	CG745119	CG745119 P037-4-G0
30	102.6	5.6	612	9	CG803681	CG803681 1118045A0
31	102.6	5.6	1289	9	AG347097	AG347097 Mus muscu
32	102.4	5.5	1491	9	CG753221	CG753221 P048-2-A0
33	102.4	5.5	1533	9	CL081968	CL081968 CH216-165
34	102.2	5.5	1378	9	AG350209	AG350209 Mus muscu
35	102.2	5.5	1758	9	CL509408	CL509408 SAIL_811
36	102	5.5	1566	9	CG757757	CG757757 P053-1-D0
37	102	5.5	1594	9	CL038406	CL038406 CH216-46A
38	101.8	5.5	1394	9	AG429511	AG429511 Mus muscu
39	101.6	5.5	1373	9	CG750869	CG750869 P045-2-E1
40	101.4	5.5	1074	8	BZ696936	BZ696936 SP_Ba009
41	101.4	5.5	1224	9	CL077121	CL077121 CH216-143
42	101.4	5.5	1461	9	CL081232	CL081232 CH216-161
43	101.4	5.5	1484	9	AG371414	AG371414 Mus muscu
44	101.4	5.5	1616	9	CL081995	CL081995 CH216-165
45	101.4	5.5	1896	9	CG753083	CG753083 P048-1-C0

ALIGNMENTS

RESULT 1
CG749499/c
LOCUS　　　　　CG749499　　　　　1348 bp　　　DNA　　　linear　　GSS 24-OCT-2003
DEFINITION　P043-4-A06.za Ppa EcoRI BAC library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION　CG749499
VERSION　　CG749499.1　GI:37970425
KEYWORDS　GSS.
SOURCE　　Pristionchus pacificus
ORGANISM　Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE　1　(bases 1 to 1348)
AUTHORS　Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE　An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL　Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE　22835951
PUBMED　12884007
COMMENT　Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 0049707160137 1
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES　Location/Qualifiers
source　1..1348
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."
ORIGIN
Query Match　6.6%; Score 121.6; DB 9; Length 1348;
Best Local Similarity　47.4%;
Matches 427; Conservative 0; Mismatches 463; Indels 10; Gaps 2;

Qy　913 TACCAGAAGCAAAACAAAAAAAATAAAACAAAGAAAAAATTTCGAGCGAAAAAATATTTG 972
　　　　　||||||||||||||||||||||||||||||||||||||||||||-||-|　||||||||||
Db　1077 TAATAAATAAAAAAATTAAAATTAAAAAAAATTTAATATATATTTTATAAAAATAAAATAA 1018

OM nucleic - nucleic search, using sw model

Run on: February 15, 2005, 09:55:11 ; Search time 1159.68 Seconds
(without alignments)
8446.087 Million cell updates/sec

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAAGGCAAATT...........AAGGACTCCGAGCCAGATCT 1660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5378673 seqs, 2950229984 residues

Total number of hits satisfying chosen parameters: 10757346

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

SUMMARIES

Result No.	Score	% Query Match Length	DB ID	Description
1	121.8	7.3	1300	18	US-10-668-749A-1	Sequence 1, Appli
2	109	6.6	6292	17	US-10-221-714A-461	Sequence 461, Appli
3	104.2	6.3	3673778	16	US-10-312-841-1	Sequence 1, Appli
4	101.4	6.1	1243	18	US-10-425-115-172717	Sequence 172717,
5	96.6	5.8	3683	18	US-10-473-126-339	Sequence 339, App
6	96	5.8	3683	18	US-10-473-126-193	Sequence 193, App
7	94.8	5.7	778	18	US-10-363-345A-2179	Sequence 2179, App
8	94.8	5.7	778	18	US-10-363-345A-2180	Sequence 2180, Ap
9	92	5.5	921	18	US-10-425-115-38710	Sequence 38710, A
10	90.6	5.5	1297	18	US-10-425-115-67687	Sequence 67687, A
11	90.2	5.4	14006	15	US-10-311-455-1931	Sequence 1931, Ap
c 12	89.8	5.4	6668	15	US-10-311-455-1670	Sequence 1670, Ap
c 13	88.2	5.3	975	18	US-10-363-345A-16099	Sequence 16099, A
c 14	88.2	5.3	975	18	US-10-363-345A-16100	Sequence 16100, A
c 15	87.8	5.3	1204	18	US-10-437-963-77858	Sequence 77858, A
c 16	87.6	5.3	7442	17	US-10-221-714A-409	Sequence 409, App
c 17	87.4	5.3	1081	18	US-10-425-115-16756	Sequence 16756, A
c 18	86.6	5.2	15548	15	US-10-311-455-2128	Sequence 2128, Ap
c 19	85	5.1	1130	18	US-10-425-115-132399	Sequence 132399,
c 20	84.6	5.1	1004	18	US-10-425-115-81106	Sequence 81106, A
c 21	84.6	5.1	1214	17	US-10-424-599-102083	Sequence 102083,
c 22	84.2	5.1	1062	18	US-10-425-115-120013	Sequence 120013,
c 23	83.8	5.0	9539	15	US-10-239-676-52	Sequence 52, Appl
c 24	83.8	5.0	9539	14	US-10-239-675-54	Sequence 54, Appl
c 25	83.6	5.0	7571	15	US-10-311-455-500	Sequence 500, App
c 26	83.6	5.0	7823	15	US-10-239-676-198	Sequence 198, App
c 27	83.6	5.0	7823	15	US-10-311-455-2034	Sequence 2034, Ap
c 28	83.6	5.0	7823	15	US-10-240-453-292	Sequence 292, App
c 29	82	4.9	7814	15	US-10-221-714A-252	Sequence 252, App
c 30	81.6	4.9	11745	15	US-10-240-453-206	Sequence 206, App
c 31	80.8	4.9	1039	18	US-10-425-115-58802	Sequence 58802, A
c 32	80	4.8	7597	18	US-10-311-455-986	Sequence 986, App
c 33	80	4.8	12237	15	US-10-311-455-2331	Sequence 2331, Ap
c 34	79.8	4.8	5276	15	US-10-311-455-123	Sequence 123, App
c 35	79	4.8	6419	15	US-10-311-455-240	Sequence 240, App
c 36	78.4	4.7	996	18	US-10-425-115-151933	Sequence 151933,
c 37	78.4	4.7	15674	15	US-10-311-455-336	Sequence 336, Appl
c 38	78.4	4.7	15674	15	US-10-311-455-2331	Sequence 2331, Ap
c 39	78.2	4.7	986	18	US-10-240-485-30	Sequence 30, Appl
c 40	78.2	4.7	9964	15	US-10-311-455-174478	Sequence 174478,
c 41	78	4.7	780	18	US-10-437-963-22707	Sequence 22707, A
c 42	78	4.7	1029	18	US-10-425-115-128448	Sequence 128448,
c 43	77.2	4.7	8056	18	US-10-473-126-386	Sequence 386, Appl
c 44	76.8	4.6	6314	16	US-10-240-452-14	Sequence 14, Appl
c 45	76.8	4.6	8056	18	US-10-473-126-386	Sequence 386, App

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
US-10-668-749A-1

; Sequence 1, Application US/10668749A
; Publication No. US20040110205A1
; GENERAL INFORMATION:
; APPLICANT: Agilent Technologies
; TITLE OF INVENTION: Methods and Systems for Nanopore Data Analysis
; FILE REFERENCE: 50112-1580
; CURRENT APPLICATION NUMBER: US/10/668,749A
; CURRENT FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.2
;
; SEQ ID NO 1
; LENGTH: 1300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-668-749A-1

Query Match 7.3%; Score 121.8; DB 18; Length 1300;
Best Local Similarity 44.3%; Pred. No. 2.8e-12;
Matches 495; Conservative 0; Mismatches 622; Indels 0; Gaps 0;

Qy 7 ACACAAGGCAAATTGAAAAAAATAGATAAAAATTTTCGCAGGTATTAAAGCCGACTTAAAC 66
|| | || |||||| |||||| |||||||||||||||
Db 1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60

Qy 67 AAATGAGTGAAGAGAAGAGAAAAAGAGAAAAAATAAATACATATATTTGAGTTAGTTAGAGAAAG 126
||| | ||| |||||||| |||||||||
Db 61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120

Qy 127 AAAAAATAAAAGAGAGACCTCGGCTTAACAGTCGAAAAACCAGAAATAATTAAAAAAGAAAGA 186
|||| ||| ||| || ||| || |||| ||| ||| ||| || |

OM nucleic - nucleic search, using sw model

Run on: February 15, 2005, 03:19:09 ; Search time 282.11 Seconds
(without alignments)
9628.226 Million cell updates/sec

Copyright (c) 1993 - 2005 Compugen Ltd.
GenCore version 5.1.6

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAAGGCAAATT.........AAGGACTCCGAGCCAGATCT 1660
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	93.8	5.7	612	4	US-09-902-540-1357	Sequence 1357, Ap
	2	82	4.9	1039	4	US-09-902-540-1280	Sequence 1280, Ap
	3	81.2	4.9	7218	1	US-08-232-463-14	Sequence 14, Appl
	4	70.2	4.2	1141	4	US-09-806-708B-22	Sequence 22, Appl
	5	69.6	4.2	19124	2	US-08-487-826B-13	Sequence 13, Appl
	6	69.2	4.2	1141	4	US-09-806-708B-22	Sequence 22, Appl
	7	67.8	4.1	5852	1	US-07-867-106-2	Sequence 2, Appli
	8	66.8	4.0	614	1	US-09-902-540-1318	Sequence 1318, Ap
c	9	64.4	3.9	1664976	4	US-09-902-540-1318	Sequence 1318, Ap
c	10	64.4	3.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	11	62.6	3.7	615	3	US-08-998-416-186	Sequence 186, App
c	12	60.6	3.7	187169	4	US-09-949-016-12776	Sequence 12776, A
c	13	60.6	3.7	191569	4	US-09-949-016-15940	Sequence 15940, A
c	14	60.4	3.6	55886	4	US-09-949-016-15129	Sequence 15129, A
c	15	60.2	3.6	1956	3	US-08-559-896B-1	Sequence 1, Appli
c	16	60.2	3.6	1956	4	US-09-351-794A-1	Sequence 1, Appli
c	17	59.6	3.6	640681	4	US-09-790-988-1	Sequence 1, Appli
c	18	59.4	3.6	205044	4	US-09-949-016-15851	Sequence 15851, A
c	19	59.4	3.6	205044	4	US-09-949-016-15852	Sequence 15852, A
c	20	59.4	3.6	205044	4	US-09-949-016-15853	Sequence 15853, A
c	21	59.4	3.6	223471	4	US-09-949-016-12387	Sequence 12387, A
c	22	59.4	3.6	223471	4	US-09-949-016-12724	Sequence 12724, A
c	23	59.4	3.6	223471	4	US-09-949-016-12725	Sequence 12725, A
c	24	59	3.6	6243	2	US-09-056-075-1	Sequence 5, Appli
c	25	58.8	3.5	10640	4	US-09-417-485D-5	Sequence 5, Appl
c	26	58.6	3.5	396	4	US-09-640-173-53	Sequence 53, Appl
c	27	58.6	3.5	396	4	US-09-713-550-53	Sequence 53, Appl
c	28	58.6	3.5	396	4	US-09-825-294-53	Sequence 53, Appl
c	29	58.6	3.5	396	4	US-09-970-966-53	Sequence 53, Appl
c	30	57.8	3.5	250352	4	US-09-949-016-14724	Sequence 14724, A
c	31	57.6	3.5	837	3	US-08-998-416-288	Sequence 288, App
c	32	57.2	3.4	601	4	US-09-949-016-30531	Sequence 30531, A
c	33	57.2	3.4	601	4	US-09-949-016-37150	Sequence 37150, A
c	34	57.2	3.4	601	4	US-09-949-016-37164	Sequence 37164, A
c	35	57.2	3.4	601	4	US-09-949-016-145868	Sequence 145868,
c	36	57.2	3.4	601	4	US-09-949-016-146136	Sequence 146136,
c	37	57	3.4	601	1	US-09-949-016-146404	Sequence 146404,
c	38	56.8	3.4	240	1	US-08-628-417-6	Sequence 6, Appli
c	39	56.8	3.4	74096	4	US-09-949-016-11785	Sequence 11785, A
c	40	56.6	3.4	74097	4	US-09-949-016-16239	Sequence 16239, A
c	41	56.6	3.4	2570	2	US-09-056-075-2	Sequence 2, Appli
c	42	55.8	3.4	636	2	US-08-998-416-1137	Sequence 1137, Ap
c	43	55.4	3.3	356	2	US-08-520-678A-22	Sequence 22, Appl
c	44	55.4	3.3	356	1	US-08-897-126-22	Sequence 22, Appl
c	45	55.4	3.3	2394	4	US-09-800-729-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-902-540-1357
Sequence 1357, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(158649)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1357
LENGTH: 612
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(612)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match   5.7%; Score 93.8; DB 4; Length 612;
Best Local Similarity 47.1%; Pred. No. 2.4e-11;
Matches   284; Conservative   0; Mismatches 319; Indels   0; Gaps   0;

```
Qy   16 AAATTGAAAAAATAGATAAAAATTTTCGCAGGTATTAAAGCCGACTTAAAACAAATGAGTG   75
          |||   |||||| |||| |   |     ||    |||       |||      |||| |
Db    1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA   60

Qy   76 AAGAGAAAAGAAAAAAAAAATAAAATACATATTTGAGTGAAAAGAGAAAGAGAAAAAATAA  135
          |||| |||||||||||||| |||||  | |  ||| |  |||  |||| |||||| |||
Db   61 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  120

Qy  136 AAGAAGACCTCGGCTTAACAGTCGAAAAACCAGAAATAATTAAAAAGAAAAGAGACTGTGAT  195
          |||| ||||  |   |||   |||||    ||||  | |  ||||||||||| ||||| ||
Db  121 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  180

Qy  196 TTTTAATGGAAATCGTCGAGGAAAAAGAAAATTTTAATTTCATTTTCGAGGGATTAATTTG  255
          ||||  |||   ||     ||||||  |||    |||  |||  ||||   |||  |||
Db  181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  240

Qy  256 TTGTAAGTTGATGAGAAAAATCTAGATAAAAAAATGCAGATCAAAAAGTGTGTCAATTTGACA  315
          |||  ||    ||  |||||  |   ||||||    ||   |||||| ||||| ||||||||
Db  241 AAAAAATAAAAAAAAAAATTAAAAAAAATAAAAAAAAATAAAAAAAAATGTGTCAATTTGACA  300
```

```
OM nucleic - nucleic search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:             February 15, 2005, 00:27:38 ; Search time 902.184 Seconds
                    (without alignments)
                    10892.200 Million cell updates/sec

Title:              US-08-876-132-1
Perfect score:      1660
Sequence:           1 AGATCTACACAAGGCAAATT..........AAGGACTCCGAGCCAGATCT 1660

Scoring table:      IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0
Searched:           4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:         8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       1: N_Geneseq_16dec04:*
                 2: geneseqn1980s:*
                 3: geneseqn1990s:*
                 3: geneseqn2000s:*
                 4: geneseqn2001as:*
                 5: geneseqn2001bs:*
                 6: geneseqn2002as:*
                 7: geneseqn2002bs:*
                 8: geneseqn2003as:*
                 9: geneseqn2003bs:*
                10: geneseqn2003cs:*
                11: geneseqn2003ds:*
                12: geneseqn2004as:*
                13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES

Result                    Query
No.     Score     Match  Length  DB  ID          Description
                    %
  1     1657    99.8    1660   2  AAV74275    Aav74275 Pantoea c
  2    121.8     7.3    1300  12  ADP85917    Adp85917 Synthetic
c 3     109      6.6    6292   4  AAS46735    Aas46735 Tumour su
c 4    107.6     6.5    1000  12  ADQ62833    Adq62833 Homopoly-
  5    107.6     6.5    1000  12  ADQ62832    Adq62832 Homopoly-
c 6     96.6      5.8    3683   8  ABZ10199    Abz10199 Haematopo
c 7      96       5.8    3683   8  ABZ10053    Abz10053 Haematopo
c 8     94.8      5.7     778   6  ABQ15588    Abq15588 Oligonucl
  9     94.8      5.7     778   6  ABQ15589    Abq15589 Oligonucl
 10     93.6      5.6    6644   2  AAX33181    Aax33181 Base sequ
 11     93.6      5.6    7372   2  AAX33182    Aax33182 Base sequ
 12     93.6      5.6    7797   2  AAX33180    Aax33180 Cowpox vi
 13     93.6      5.6    7996   2  AAX33184    Aax33184 Base sequ
c 14    90.4      5.4    8222   8  ACF62816    Acf62816 Colon can
 15     90.4      5.4   11222  10  ADB54318    Adb54318 Pretreate
 16     90.2      5.4   14006   6  ABL33958    Abl33958 Human imm
 17     89.8      5.4    6668   6  ABL33697    Abl33697 Human imm
c 18     88.8      5.3    8222   8  ACF62794    Acf62794 Colon can
c 19     88.8      5.3   11222  10  ADB54190    Adb54190 Pretreate
c 20     88.2      5.3     975   6  ABQ29508    Abq29508 Oligonucl
c 21     88.2      5.3     975   6  ABQ29509    Abq29509 Oligonucl
c 22     87.6      5.3    7442   4  AAS46686    Aas46686 Tumour su
c 23     86.6      5.2   15548   6  ABL34155    Abl34155 Human imm
c 24     83.8      5.0    9539   4  AAS45347    Aas45347 Chemical1
c 25     83.8      5.0    9539   4  AAS45347    Aas45347 Chemical1
c 26     83.6      5.0    9539   6  ABK28180    Abk28180 DNA trans
c 27     83.6      5.0    7571   6  ABL32527    Abl32527 Human imm
c 28     83.6      4.5   7823   4  AAS45490    Aas45490 Chemical1
c 29     83.6      5.0    7823   6  ABL34061    Abl34061 Human imm
c 30     83.6      5.0    7823   6  ABK31493    Abk31493 Signal tr
c 31     82        4.9   7814   4  AAS46530    Aas46530 Tumour su
c 32     81.6      4.9   11745   5  ABK28332    Abk28332 DNA trans
c 33     80        4.8   7597   6  ABL33013    Abl33013 Human imm
c 34     80        4.8   12237   6  ABL34358    Abl34358 Human imm
c 35     79.8      4.8   5276   6  ABL32150    Abl32150 Human imm
c 36     79        4.8   6419   6  ABL32267    Abl32267 Human imm
c 37     78.4      4.7   2943   3  AAA70229    Aaa70229 Plasmodiu
c 38     78.4      4.7   15674   6  ABL32363    Abl32363 Human imm
c 39     78.4      4.7   15674   6  ABL34477    Abl34477 Human met
c 40     78.4      4.7   15674   6  ABL70514    Abl70514 Chemical1
c 41     78.4      4.7   15674   7  ADS99738    Ads99738 Complemen
c 42     78.2      4.7   9964   6  ABL32098    Abl32098 Human imm
c 43     77.4      4.7   5493  13  ADS89675    Ads89675 Oligonucl
c 44     77.2      4.7   8056   8  ABZ10246    Abz10246 Haematopo
c 45     76.8      4.6   6314   6  ABL54314    Abl54314 Chemical1

                              ALIGNMENTS

RESULT 1
AAV74275
ID   AAV74275 standard; DNA; 1660 BP.
XX
AC   AAV74275;
XX
DT   16-JUN-1999  (first entry)
XX
DE   Pantoea citrea cryptic plasmid pS.
XX
KW   Cryptic plasmid; pS; higher temperature; growth; elimination;
KW   mobilization; ds.
XX
OS   Pantoea citrea.
XX
FH   Key             Location/Qualifiers
FT   CDS             480..1595
FT                   /*tag= a
FT                   /note= "largest open reading frame"
XX
PN   WO9859054-A1.
XX
PD   30-DEC-1998.
XX
PF   22-JUN-1998;  98WO-US012945.
XX
PR   23-JUN-1997;  97US-00876132.
XX
PA   (GEMV ) GENENCOR INT INC.
XX
PI   Fowler T, Causey SC;
XX
DR   WPI; 1999-105625/09.
DR   P-PSDB; AAW90181.
XX
PT   Method for preparing an improved Enterobacteriaceae strain - useful for
PT   improving bacterial fermentation strains of the family
PT   Enterobacteriaceae.
XX
PS   Claim 18; Fig 1A-C; 30pp; English.
XX
CC   The sequence is that of one half of the cryptic plasmid pS from Pantoea
CC   citrea. It can be used as part of a method of eliminating the cryptic
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2005, 00:33:38 ; Search time 7168.72 Seconds
                                            (without alignments)
                                            11220.366 Million cell updates/sec

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAAGGCAAATT.........AAGGACTCCGAGCCAGATCT 1660

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422760795 residues

Total number of hits satisfying chosen parameters: 9416466
Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :
       GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
       10: gb_ro:*
       11: gb_sts:*
       12: gb_sy:*
       13: gb_un:*
       14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB ID	Description
	1	1308.6	78.8	3661	8	AF128889	Pantoea c
c	2	128	7.7	1328	8	AJ840631	Arabidops
c	3	120.4	7.3	170627	2	AC125567	Rattus no
c	4	117.6	7.1	66993	2	AC138074	Homo sapi
c	5	117.6	7.1	1142	8	AJ840525	Arabidops
c	6	112.4	6.8	93791	2	AC138073	Homo sapi
c	7	109	6.6	6292	6	AX251493	Sequence
c	8	106.2	6.4	108908	6	AX251493	Homo sapi
c	9	104.2	6.3	349980	6	AX344555	Sequence
c	10	104	6.3	286208	2	AL034560	Plasmodiu
c	11	101.6	6.1	131682	2	AC117140	Rattus no
c	12	101.6	6.1	231912	2	AL672277	Human DNA
c	13	99.6	6.0	174384	2	AC009524	Mus muscu
c	14	98.4	5.9	110000	2	AC009524	Homo sapi
c	15	98.4	5.9	130540	2	AC116305_3	Continuation (4 of
c	16	96.6	5.8	3683	6	AX598999	Mus muscu
c	17	96.6	5.8	14867	3	AE001398	Plasmodiu
c	18	96.6	5.8	84563	5	AC096885	Danio rer
c	19	96	5.8	3683	6	AX598853	Sequence
	20	95.4	5.7	83435	5	BX640469	Zebrafish
	21	93.8	5.7	175544	2	AC117342	Rattus no
	22	93.8	5.7	289973	2	AC135678	Rattus no
c	23	93.8	5.7	340089	2	AC104922	Mus muscu
	24	93.6	5.6	6644	6	E23356	Virus vecto
	25	93.6	5.6	7372	6	E23357	Virus vecto
	26	93.6	5.6	7797	6	E23355	Virus vecto
	27	93.6	5.6	7996	6	E23359	Virus vecto
	28	93.6	5.6	1056082	3	AC116957_3	Continuation (4 of
	29	93.4	5.6	1197	8	AJ840663	Arabidops
	30	91.8	5.6	67970	3	PFMAL1P3	Plasmodiu
c	31	91.8	5.5	160667	2	AC151623	Bos tauru
c	32	91.2	5.5	347050	2	PFA029351	Plasmodiu
c	33	90.8	5.5	257703	3	CEY111B2A	Caenorhab
c	34	90.6	5.5	131026	3	BX004843	Zebrafish
c	35	90.4	5.4	8222	6	AX705396	Sequence
c	36	90.4	5.4	11222	6	AX822482	Sequence
c	37	90.4	5.4	11222	6	AX826122	Sequence
c	38	90.2	5.4	197951	2	AX346860	Sequence
c	39	90.2	5.4	14951	6	CR392343	Danio rer
c	40	90.2	5.4	206038	2	CR847802	Danio rer
c	41	90.2	5.4	239339	2	CR847802	Danio rer
c	42	90	5.4	206038	2	CEY111B2_0	Caenorhabdi
c	43	90	5.4	300695	2	AC079431	Mus muscu
c	44	89.8	5.4	6668	6	AX346599	Sequence
c	45	89.8	5.4	110743	5	BX649389	Zebrafish

ALIGNMENTS

RESULT 1
AF128889
LOCUS   AF128889   3661 bp   DNA   circular BCT 22-MAR-1999
DEFINITION Pantoea citrea strain ATCC31623 plasmid pPZG500, complete plasmid
            sequence.
ACCESSION   AF128889
VERSION   AF128889.1 GI:4457229
KEYWORDS   .
SOURCE   Pantoea citrea
  ORGANISM Pantoea citrea
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Pantoea.
REFERENCE   1 (bases 1 to 3661)
  AUTHORS   Bilic,M. and Delic,V.
  TITLE   Isolation and characterization of a cryptic plasmid from Erwinia
            citreus ATCC 31623
  JOURNAL   J. Appl. Microbiol. 83 (4), 485-492 (1997)
  MEDLINE   98012514
   PUBMED   9351229
REFERENCE   2 (bases 1 to 3661)
  AUTHORS   Bilic,M. and Delic,V.
  TITLE   Sequence Analysis and Functional Characterization of Replication
            Region in Plasmid pPZG500 from Pantoea citrea ATCC 31623
  JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 3661)
  AUTHORS   Bilic Nezic,M. and Delic,V.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-FEB-1999) Biotechnology, PLIVA d.d., Research
            Institute, Prilaz baruna Filipovica 25, Zagreb 10 000, Croatia
FEATURES   Location/Qualifiers
     source   1..3661
                     /organism="Pantoea citrea"
                     /mol_type="genomic DNA"
                     /strain="ATCC31623"
                     /db_xref="ATCC:31623"
                     /db_xref="taxon:53336"
                     /plasmid="pPZG500"
     misc_RNA   complement(join(3188..3661,1..177))
                     /note="RNAII"
     CDS   99..317
                     /note="Orf1"
                     /codon_start=1

```
OM nucleic - nucleic search, using sw model

                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:             February 15, 2005, 00:53:24 ; Search time 5561.26 Seconds
                                               (without alignments)
                                               11361.937 Million cell updates/sec

Title:              US-08-876-132-1
Perfect score:      1660
Sequence:           1 AGATCTACACAAGGCAAATT..........AAGGACTCCGAGCCAGATCT 1660

Scoring table:      IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0

Searched:           34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters:       68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :          EST:*
                 1: gb_est1:*
                 2: gb_est2:*
                 3: gb_htc:*
                 4: gb_est3:*
                 5: gb_est4:*
                 6: gb_est5:*
                 7: gb_est6:*
                 8: gb_gss1:*
                 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                                     SUMMARIES

Result              %
No.     Score   Query Match Length DB ID      Description

  1     147.2   8.9         1594   9  CL078613    CL078613 CH216-151
c 2     143     8.6         1407   9  AJ592026    AJ592026 Arabidops
c 3     141.2   8.5         1599   9  CL083840    CL083840 CH216-167
c 4     140.8   8.5         1632   9  CL082569    CL082569 CH216-167
c 5     140     8.4         1842   9  CL044158    CL044158 CH216-59P
c 6     138.2   8.3         1641   9  CG753221    CG753221 P048-2-A0
c 7     137.8   8.3         1641   9  CL078595    CL078595 CH216-151
c 8     135.6   8.2         1493   9  CL078589    CL078589 CH216-151
c 9     135.4   8.2         1313   9  CG751144    CG751144 P045-3-E0
c 10    135.2   8.1         1376   9  CG747831    CG747831 P041-3-B0
c 11    134.8   8.1         1536   9  CL078538    CL078538 CH216-151
c 12    134.6   8.1         1434   9  AJ592058    AJ592058 Arabidops
c 13    133.8   8.1         1311   9  AG346751    AG346751 Mus muscu
c 14    133.6   8.0         1453   9  AJ591978    AJ591978 Arabidops
c 15    133.2   8.0         1334   9  AG341525    AG341525 Mus muscu
c 16    133.2   8.0         1616   9  CL081995    CL081995 CH216-165
c 17    133     8.0         1482   9  AG332292    AG332292 Mus muscu
c 18    132.4   8.0         1594   9  CL038406    CL038406 CH216-46A
c 19    132     8.0         1519   9  AG386893    AG386893 Mus muscu
c 20    131.8   7.9         1566   9  CG757757    CG757757 P053-1-D0
c 21    131.6   7.9         1784   9  CL081992    CL081992 CH216-165
c 22    131.4   7.9         1276   9  CG754010    CG754010 P049-2-C0
c 23    131.4   7.9         1981   9  CL082000    CL082000 CH216-165
c 24    130.8   7.9         1259   9  CR004722    CR004722 Forward 8
c 25    130.8   7.9         1512   9  CL082685    CL082685 CH216-169
c 26    130.6   7.9         1531   9  CG748014    CG748014 P041-4-B0
c 27    130.4   7.9         1428   9  CL081978    CL081978 CH216-165
c 28    130.2   7.8         1608   9  CL118721    CL118721 ISB1-72J8
c 29    130     7.8         1472   9  AG350118    AG350118 Mus muscu
c 30    129.8   7.8         1811   9  CG753732    CG753732 P048-4-G0
c 31    129.4   7.8         1615   9  AG311072    AG311072 Mus muscu
c 32    129.2   7.8         1434   8  CC187638    CC187638 Arabidops
c 33    128.8   7.8         1288   9  AG347321    AG347321 Mus muscu
c 34    128.6   7.7         1727   9  CL114085    CL114085 ISB1-59J1
c 35    128.2   7.7         1805   9  CL080711    CL080711 CH216-159
c 36    128.2   7.7         1165   9  CR007727    CR007727 Reverse s
c 37    128     7.7         1507   9  AG346189    AG346189 Mus muscu
c 38    127.8   7.7         1228   9  AJ592059    AJ592059 Arabidops
c 39    127.6   7.7         1361   9  CG744327    CG744327 P036-4-E0
c 40    127     7.7         1533   9  CL081968    CL081968 CH216-165
c 41    126.8   7.7         1385   9  CL076799    CL076799 CH216-140
c 42    126.8   7.6         1329   9  CL078715    CL078715 CH216-152
c 43    126.8   7.6         1459   9  CL040025    CL040025 CH216-49A
c 44    126.6   7.6         1843   9  AG435185    AG435185 Mus muscu
c 45    126.6   7.6         1457   9  CL082658    CL082658 CH216-169
```

                                     ALIGNMENTS

```
RESULT 1
CL078613

LOCUS       CL078613             1594 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-151M17_Sp5.1 CH216 Xenopus tropicalis genomic clone
            CH216-151M17, genomic survey sequence.
ACCESSION   CL078613
VERSION     CL078613.1  GI:40534526
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1594)
  AUTHORS   Krementzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp5 atctgccgttcgatcct
            Class: BAC ends
            High quality sequence start: 557
            High quality sequence stop: 634.
FEATURES             Location/Qualifiers
     source          1..1594
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-151M17"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

Query Match  8.9%;  Score 147.2;  DB 9;  Length 1594;
Best Local Similarity  44.5%;  Pred. No. 5.2e-17;
Matches 500; Conservative  0; Mismatches 623; Indels  0; Gaps  0;

Qy          1 AGATCTACACAAGGCAAATTGAAAAATTGAAAAAATTGATAAAAATTTTCGCGAGTATTAAAGCCGACT 60
              || ||| ||| ||  |  ||| || ||  |  ||| || |  || || | | | |  |  || |  |  |
```